Support Vector Machines: Maximum Margin Classifiers

Nipun Batra

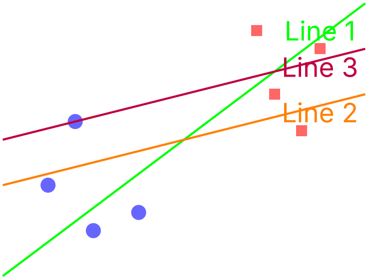
IIT Gandhinagar

July 30, 2025

Outline

The Problem with Many Decision Boundaries

Question: For linearly separable data, which line should we choose?



All lines separate the data perfectly! Which is best for new, unseen data?

Pop Quiz: Decision Boundary Selection

Quick Quiz 1

Among all possible separating lines, which criterion should we use to choose the best one?

a) The line closest to the data points

Answer: b) Maximum margin! The line farthest from both classes generalizes better.

Pop Quiz: Decision Boundary Selection

Quick Quiz 1

Among all possible separating lines, which criterion should we use to choose the best one?

a) The line closest to the data points
b) The line that maximizes distance to closest points

Answer: b) Maximum margin! The line farthest from both classes generalizes better.

Pop Quiz: Decision Boundary Selection

Quick Quiz 1

Among all possible separating lines, which criterion should we use to choose the best one?

a) The line closest to the data points
b) The line that maximizes distance to closest points
c) Any line works equally well

Answer: b) Maximum margin! The line farthest from both classes generalizes better.

SVM's Big Idea: Maximum Margin

SVM Principle

Among all separating hyperplanes, choose the one that **maximizes the margin** to the nearest data points.

SVM's Big Idea: Maximum Margin

SVM Principle

Among all separating hyperplanes, choose the one that **maximizes the margin** to the nearest data points.

SVM's Big Idea: Maximum Margin

SVM Principle

Among all separating hyperplanes, choose the one that **maximizes the margin** to the nearest data points.

Key concepts:

- **Margin:** Distance from decision boundary to closest points

SVM's Big Idea: Maximum Margin

SVM Principle

Among all separating hyperplanes, choose the one that **maximizes the margin** to the nearest data points.

Key concepts:

- **Margin:** Distance from decision boundary to closest points
- **Support Vectors:** The closest points that define the margin

SVM's Big Idea: Maximum Margin

SVM Principle

Among all separating hyperplanes, choose the one that **maximizes the margin** to the nearest data points.

Key concepts:

- **Margin:** Distance from decision boundary to closest points
- **Support Vectors:** The closest points that define the margin
- **Maximum Margin:** Optimal decision boundary

SVM's Big Idea: Maximum Margin

SVM Principle

Among all separating hyperplanes, choose the one that **maximizes the margin** to the nearest data points.

Key concepts:

- **Margin:** Distance from decision boundary to closest points
- **Support Vectors:** The closest points that define the margin
- **Maximum Margin:** Optimal decision boundary

SVM's Big Idea: Maximum Margin

SVM Principle

Among all separating hyperplanes, choose the one that **maximizes the margin** to the nearest data points.

Key concepts:

- **Margin:** Distance from decision boundary to closest points
- **Support Vectors:** The closest points that define the margin
- **Maximum Margin:** Optimal decision boundary

Intuition

Larger margin = More confident predictions = Better generalization to new data

Mathematical Formulation: Distance Between Hyperplanes

Given two parallel hyperplanes:

- $\mathbf{w} \cdot \mathbf{x} + b_1 = 0$

Mathematical Formulation: Distance Between Hyperplanes

Given two parallel hyperplanes:

- $\mathbf{w} \cdot \mathbf{x} + b_1 = 0$
- $\mathbf{w} \cdot \mathbf{x} + b_2 = 0$

Mathematical Formulation: Distance Between Hyperplanes

Given two parallel hyperplanes:

- $\mathbf{w} \cdot \mathbf{x} + b_1 = 0$
- $\mathbf{w} \cdot \mathbf{x} + b_2 = 0$

Mathematical Formulation: Distance Between Hyperplanes

Given two parallel hyperplanes:

- $\mathbf{w} \cdot \mathbf{x} + b_1 = 0$
- $\mathbf{w} \cdot \mathbf{x} + b_2 = 0$

Distance formula: $D = \frac{|b_1 - b_2|}{\|\mathbf{w}\|}$

Mathematical Formulation: Distance Between Hyperplanes

Given two parallel hyperplanes:

- $\mathbf{w} \cdot \mathbf{x} + b_1 = 0$
- $\mathbf{w} \cdot \mathbf{x} + b_2 = 0$

Distance formula: $D = \frac{|b_1 - b_2|}{\|\mathbf{w}\|}$

Derivation Intuition

Mathematical Formulation: Distance Between Hyperplanes

Given two parallel hyperplanes:

- $\mathbf{w} \cdot \mathbf{x} + b_1 = 0$
- $\mathbf{w} \cdot \mathbf{x} + b_2 = 0$

Distance formula: $D = \frac{|b_1 - b_2|}{||\mathbf{w}||}$

Derivation Intuition

- Find point $\mathbf{x}_1$ on first hyperplane

Mathematical Formulation: Distance Between Hyperplanes

Given two parallel hyperplanes:

- $\mathbf{w} \cdot \mathbf{x} + b_1 = 0$
- $\mathbf{w} \cdot \mathbf{x} + b_2 = 0$

Distance formula: $D = \frac{|b_1 - b_2|}{||\mathbf{w}||}$

Derivation Intuition

- Find point $\mathbf{x}_1$ on first hyperplane
- Move perpendicular distance t in direction $\mathbf{w}$: $\mathbf{x}_2 = \mathbf{x}_1 + t\frac{\mathbf{w}}{||\mathbf{w}||}$

Mathematical Formulation: Distance Between Hyperplanes

Given two parallel hyperplanes:

- $\mathbf{w} \cdot \mathbf{x} + b_1 = 0$
- $\mathbf{w} \cdot \mathbf{x} + b_2 = 0$

Distance formula: $D = \frac{|b_1 - b_2|}{\|\mathbf{w}\|}$

Derivation Intuition

- Find point $\mathbf{x}_1$ on first hyperplane
- Move perpendicular distance t in direction $\mathbf{w}$: $\mathbf{x}_2 = \mathbf{x}_1 + t\frac{\mathbf{w}}{\|\mathbf{w}\|}$
- Substitute into second hyperplane equation

Mathematical Formulation: Distance Between Hyperplanes

Given two parallel hyperplanes:

- $\mathbf{w} \cdot \mathbf{x} + b_1 = 0$
- $\mathbf{w} \cdot \mathbf{x} + b_2 = 0$

Distance formula: $D = \frac{|b_1 - b_2|}{\|\mathbf{w}\|}$

Derivation Intuition

- Find point $\mathbf{x}_1$ on first hyperplane
- Move perpendicular distance t in direction $\mathbf{w}$: $\mathbf{x}_2 = \mathbf{x}_1 + t \frac{\mathbf{w}}{\|\mathbf{w}\|}$
- Substitute into second hyperplane equation
- Solve for t to get distance

Pop Quiz: Hyperplane Distance

Quick Quiz 2

What is the distance between hyperplanes $2x + y - 3 = 0$ and $2x + y + 1 = 0$?

a) 1 unit

Answer: c) $D = \frac{|(-3)-1|}{\sqrt{2^2+1^2}} = \frac{4}{\sqrt{5}}$ units

Pop Quiz: Hyperplane Distance

Quick Quiz 2

What is the distance between hyperplanes $2x + y - 3 = 0$ and $2x + y + 1 = 0$?

a) 1 unit
b) 2 units

Answer: c) $D = \frac{|(-3)-1|}{\sqrt{2^2+1^2}} = \frac{4}{\sqrt{5}}$ units

Pop Quiz: Hyperplane Distance

Quick Quiz 2

What is the distance between hyperplanes $2x+y-3=0$ and $2x+y+1=0$?

a) 1 unit
b) 2 units
c) $\frac{4}{\sqrt{5}}$ units

Answer: c) $D = \frac{|(-3)-1|}{\sqrt{2^2+1^2}} = \frac{4}{\sqrt{5}}$ units

The SVM Optimization Problem

Goal: Find hyperplane that maximizes margin

Hard-Margin SVM Formulation

$$\text{Minimize: } \frac{1}{2}\|\mathbf{w}\|^2 \tag{1}$$

$$\text{Subject to: } y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \quad \forall i \tag{2}$$

The SVM Optimization Problem

Goal: Find hyperplane that maximizes margin

Hard-Margin SVM Formulation

$$\text{Minimize: } \frac{1}{2}\|\mathbf{w}\|^2 \quad (1)$$

$$\text{Subject to: } y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \quad \forall i \quad (2)$$

The SVM Optimization Problem

Goal: Find hyperplane that maximizes margin

Hard-Margin SVM Formulation

$$\text{Minimize: } \frac{1}{2}\|\mathbf{w}\|^2 \tag{1}$$

$$\text{Subject to: } y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \quad \forall i \tag{2}$$

Why this formulation?

- Minimizing $\|\mathbf{w}\|^2 \Rightarrow$ Maximizing margin $\frac{2}{\|\mathbf{w}\|}$

The SVM Optimization Problem

Goal: Find hyperplane that maximizes margin

Hard-Margin SVM Formulation

$$\text{Minimize: } \frac{1}{2}\|\mathbf{w}\|^2 \tag{1}$$

$$\text{Subject to: } y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \quad \forall i \tag{2}$$

Why this formulation?

- Minimizing $\|\mathbf{w}\|^2 \Rightarrow$ Maximizing margin $\frac{2}{\|\mathbf{w}\|}$
- Constraint ensures all points are correctly classified

The SVM Optimization Problem

Goal: Find hyperplane that maximizes margin

Hard-Margin SVM Formulation

$$\text{Minimize: } \frac{1}{2}\|\mathbf{w}\|^2 \tag{1}$$

$$\text{Subject to: } y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \quad \forall i \tag{2}$$

Why this formulation?

- Minimizing $\|\mathbf{w}\|^2 \Rightarrow$ Maximizing margin $\frac{2}{\|\mathbf{w}\|}$
- Constraint ensures all points are correctly classified
- Support vectors lie exactly on margin boundary: $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) = 1$

Pop Quiz: SVM Constraints

Quick Quiz 3

In the SVM constraint $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1$, what does $y_i = +1$ and $\mathbf{w} \cdot \mathbf{x}_i + b = 0.5$ mean?

a) Point is correctly classified and satisfies constraint

Answer: b) Since $1 \times 0.5 = 0.5 < 1$, the constraint is violated!

Pop Quiz: SVM Constraints

Quick Quiz 3

In the SVM constraint $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1$, what does $y_i = +1$ and $\mathbf{w} \cdot \mathbf{x}_i + b = 0.5$ mean?

a) Point is correctly classified and satisfies constraint
b) Point violates the constraint

Answer: b) Since $1 \times 0.5 = 0.5 < 1$, the constraint is violated!

Pop Quiz: SVM Constraints

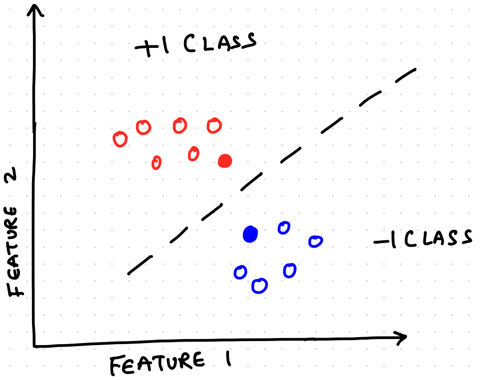
Quick Quiz 3

In the SVM constraint $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1$, what does $y_i = +1$ and $\mathbf{w} \cdot \mathbf{x}_i + b = 0.5$ mean?

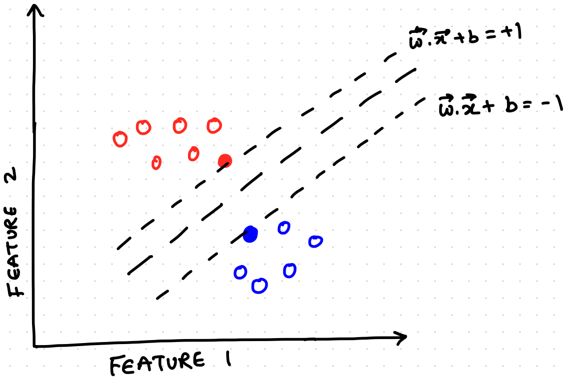
a) Point is correctly classified and satisfies constraint
b) Point violates the constraint
c) Point is a support vector

Answer: b) Since $1 \times 0.5 = 0.5 < 1$, the constraint is violated!

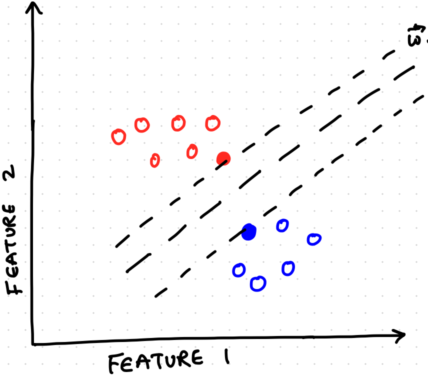
FORMULATION



FORMULATION



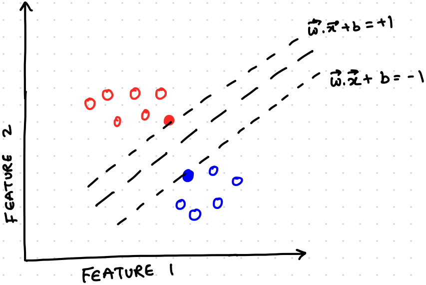
FORMULATION



$$\text{MARGIN} = \frac{(b+1)-(b-1)}{\|\vec{w}\|}$$

$$= \frac{2}{\|\vec{w}\|}$$

FORMULATION



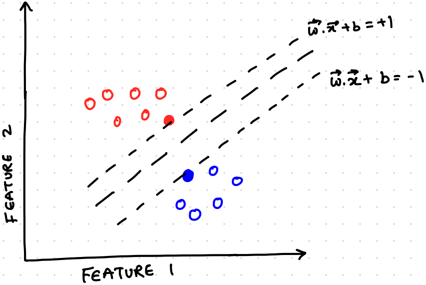
GOAL: MAXIMIZE MARGIN

$\Rightarrow$ MAXIMIZE $\dfrac{2}{||\vec{w}||}$

$\Rightarrow$ MINIMIZE $||\vec{w}||$

S.T. Correctly label points

FORMULATION



GOAL: MAXIMIZE MARGIN

$\Rightarrow$ MAXIMIZE $\dfrac{2}{\|\vec{w}\|}$

$\Rightarrow$ MINIMIZE $\|\vec{w}\|$

S.T. correctly label points

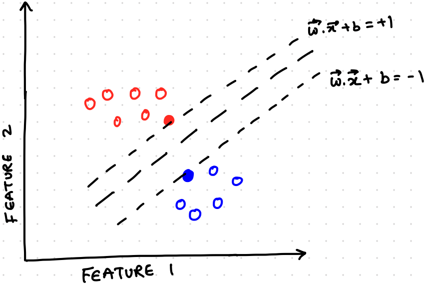
i.e. if $y_i = -1$

$\vec{w}.\vec{x}+b \leq -1$

if $y_i = +1$

$\vec{w}.\vec{x}+b \geq 1$

FORMULATION



GOAL: MAXIMIZE MARGIN

$\Rightarrow$ MAXIMIZE $\dfrac{2}{\|\vec{w}\|}$

$\Rightarrow$ $\boxed{\text{MINIMIZE } \|\vec{w}\|}$

S.T. correctly label points

i.e. if $y_i = -1$

$\vec{w}.\vec{x}+b \leq -1$

if $y_i = +1$

$\vec{w}.\vec{x}+b \geq 1$

$$\boxed{y_i(\vec{w}.\vec{x}+b) \geq 1}$$

Q) What is $\|\mathbf{w}\|$?

Q) What is $\|\mathbf{w}\|$?

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

$$\begin{aligned} \|\mathbf{w}\| &= \sqrt{\mathbf{w}^\top \mathbf{w}} \\ &= \sqrt{\begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}} \end{aligned}$$

EXAMPLE (IN 1D)